

2. Meteorological Data Processing with Calmet

The NDDH processed five years of meteorological data (1990-1994) to use with Calpuff. Raw meteorological data was derived from National Weather Service, Federal Aviation Administration, U.S. Military, and Environment Canada observations.

2.1 Calmet Grid

The Calmet grid utilized by the NDDH was established in a generic sense, and was not customized for the Milton R. Young analysis. In establishing grid size, the primary goal was to provide a modeling domain which would encompass new or existing emission sources located up to 250 km from any North Dakota Class I area. To accommodate NDDH administrator concerns regarding previous air quality issues with the state of Minnesota, the grid was then extended eastward to include the western part of Minnesota. The dimensions of the resultant grid are 800 km east-west and 540 km north-south. The extent of the NDDH Calmet grid is illustrated in Figure 2-1.

Because of the size of the NDDH Calmet grid, some compromise was necessary in the selection of grid cell size. To keep computer disk storage and model execution time requirements practical, grid cell size was set to 20 km. Even at this resolution, a single year of Calmet-processed meteorological data requires almost one gigabyte of disk space. Given the gently rolling nature of terrain, relatively uniform land-use characteristics, and the general lack of large terrain features or water bodies large enough to cause persistent, strong local-scale flows, the NDDH believes this compromise should not adversely affect model performance. Grid cell size is also depicted in Figure 2-1.

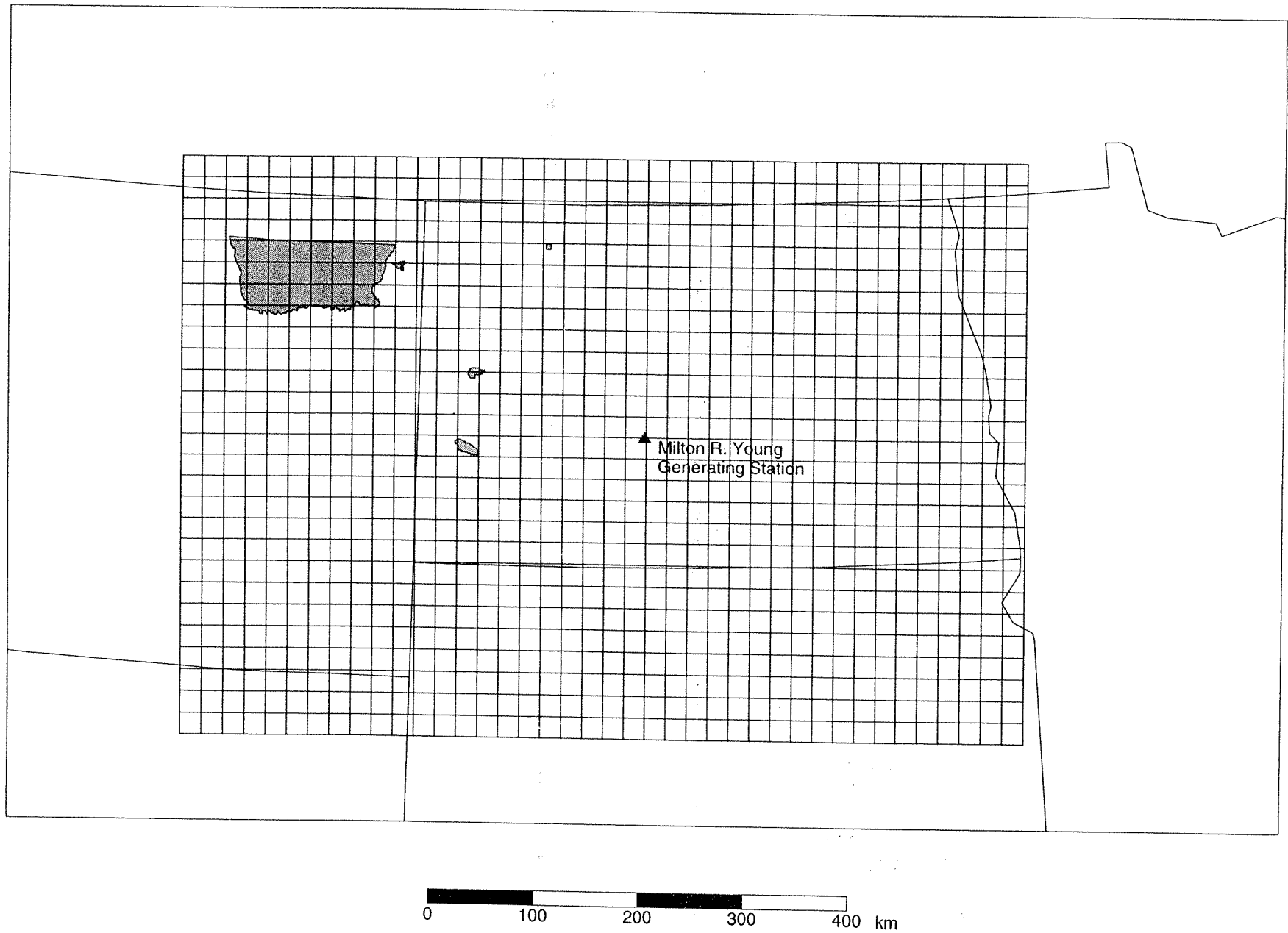
In the vertical, the NDDH Calmet grid is defined by eight vertical layers. Cell face heights are set at 22, 50, 100, 250, 500, 1000, 2000, and 4000 meters above ground level (AGL). The use of eight layers is somewhat arbitrary, but appears consistent with some of the examples and guidance provided with the modeling system.

Because the NDDH Calmet domain is large, the grid system, meteorological data, and geophysical data were fit to Lambert conformal mapping to account for the earth's curvature.

2.2 Surface Meteorological Data

Surface meteorological data for the five-year period 1990-1994 were obtained in TD-1440 format from the National Climatic Data Center (NCDC). Data were obtained for 25 stations (National Weather

Figure 2-1: Location of Calmet 20km Grid



Service, Federal Aviation Administration, U.S. Military, Environment Canada) located within or near the NDDH Calmet grid. Location of these stations is shown in Figure 2-2.

Some adjustments to the surface data files were required before Earth Tech programs METSCAN and SMERGE could be applied. Stations other than first-order National Weather Service (NWS) were missing opaque cloud cover for the entire five-year period. Based on a comparison of total and opaque cloud cover in the first-order NWS data sets, the NDDH developed an objective scheme to extrapolate opaque from total cloud cover. This scheme was coded into a computer program (TOT2OPQ) and applied to all surface data sets with missing opaque cloud cover.

EPA recommendations were followed to substitute for other missing data^{7,8} (i.e., ceiling height, wind, pressure, temperature, relative humidity). Substitutions were made if data elements were missing for one or two consecutive hours. Except for opaque cloud cover, substitutions were not made for longer missing periods (Calmet ignores stations with missing data). The EPA substitution scheme was coded into a computer program (SUB144) and applied to all surface data sets.

Earth Tech program METSCAN was next applied to scan each data set for missing or unreasonable values. A few very minor changes were resultantly applied. Lastly, Earth Tech program SMERGE was applied to merge individual station data sets into a single input file (SURF.DAT) compatible with Calmet.

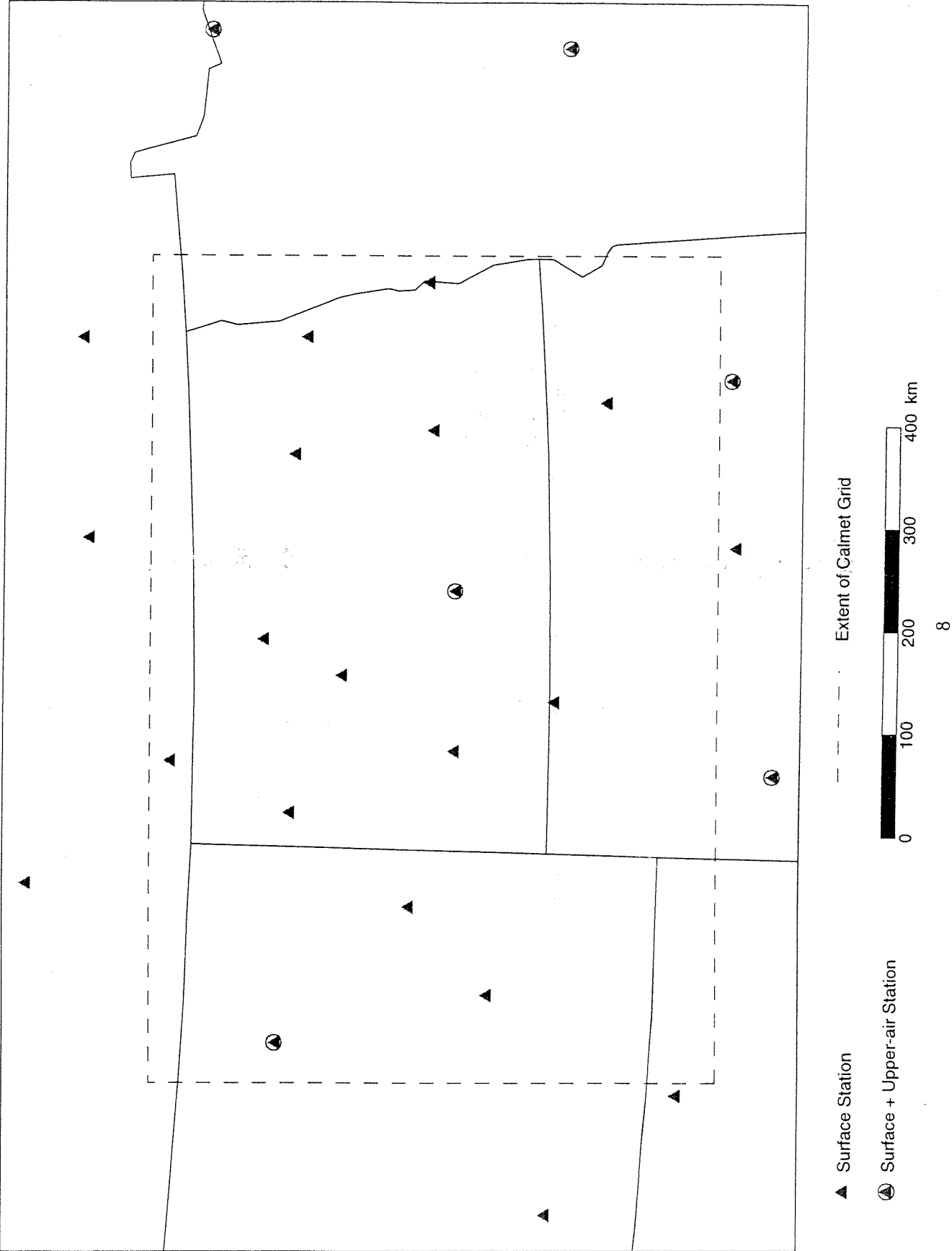
The NDDH procedure for preparation of surface meteorological data is depicted schematically in Figure 2-3.

2.3 Upper-Air Meteorological Data

Upper-air meteorological data for 1990 through 1994 were obtained in TD-6201 format from NCDC. Data were obtained for six upper-air stations located within or near the NDDH Calmet grid. The locations of these stations are shown in Figure 2-2.

Because of Calmet's fairly strict requirements on the completeness of upper-air data records and the frequency of missing upper-air data, it was desirable to automate the upper-air data processing with computer programs as much as possible. Since Earth Tech's program READ62 did not correct errors or fill in for missing data, much of the upper-air data processing was accomplished by running programs written by NDDH staff, along with a fair amount of manual file editing. The procedure consisted of preparing the TD-6201 data file for Earth Tech's program READ62, execution of a modified

Figure 2-2: Surface / Upper-air Meteorological Stations



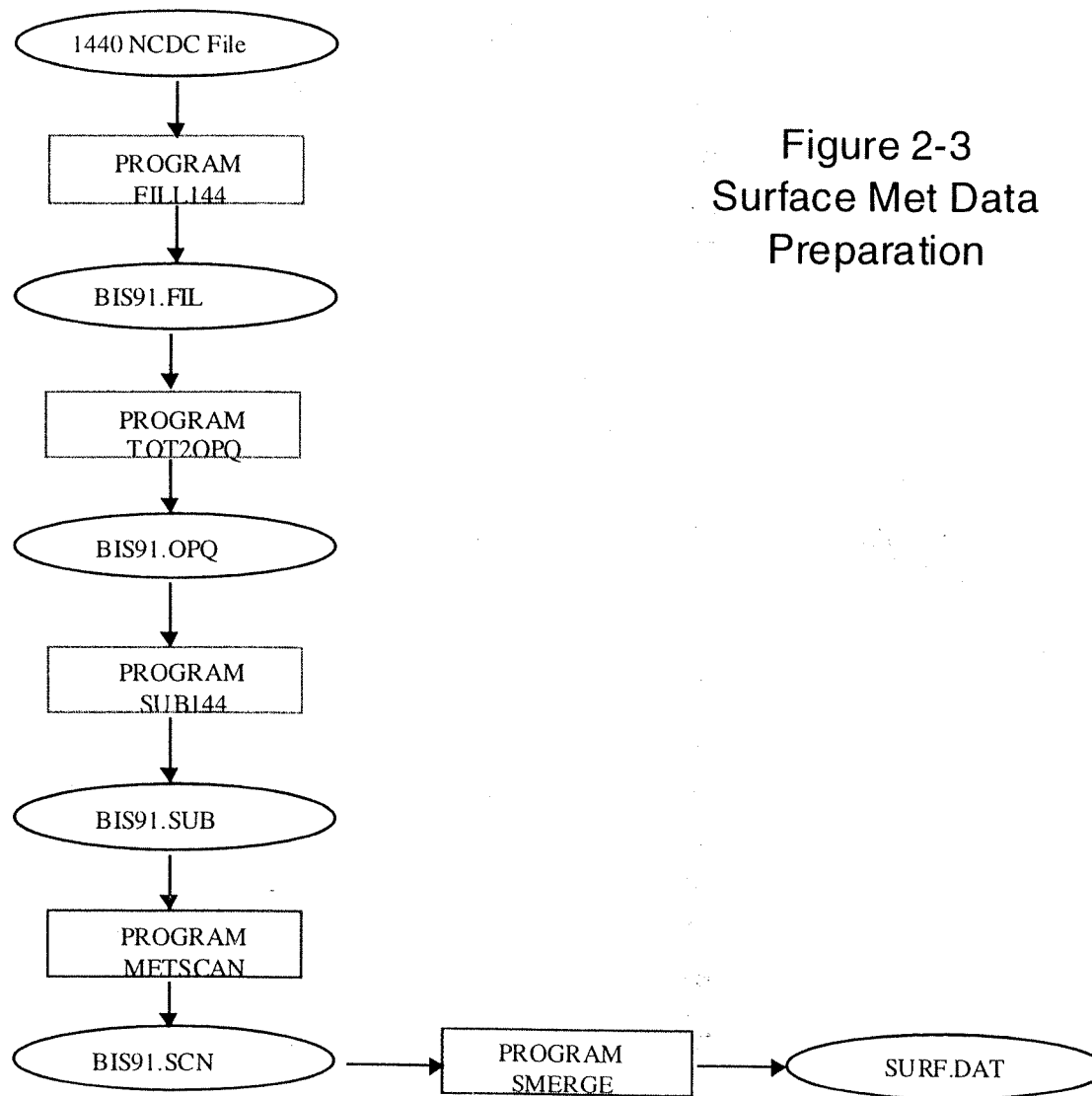


Figure 2-3
Surface Met Data
Preparation

version of READ62, execution of two NDDH programs to fill in some missing data, and some manual editing to handle more complicated problems or fill in extended missing periods.

The first NDDH upper-air data program (DEL62) input the TD-6201 sounding data for one station for one year, made the hours of observation consistent at 00Z and 12Z, and output only the first line of data (lowest 79 sounding levels) for each sounding. Then, program READ62 was executed to convert upper-air data to the format Calmet requires and output only the desired variables pressure, height (MSL), temperature, and wind at sounding levels up to the top level desired, 500 mb in this case. Calmet requires sounding data at levels up to the calculated mixing height for all hours processed. It was necessary to extract data up to 500 mb to accommodate mixing heights up to approximately 4 km AGL at stations as high as Rapid City, South Dakota (elevation 966 m MSL).

Since program READ62 only flags missing soundings but does not fill in any missing data, the rest of the data preparation and missing data substitution is left up to the user. In addition, READ62 (version available in 1996-97) created some problems that were corrected in later NDDH programs or by modifying READ62. READ62 rejected some surface data records because they were coded as having been substituted for missing data by NCDC. Since surface data must be complete and the NCDC substitutions looked reasonable, READ62 was modified to accept this type of data. Also, READ62 overlooked some occurrences of two consecutive missing soundings and did not code missing temperatures correctly for Calmet input, which were remedied in a later NDDH program.

Two NDDH programs (FIXR62 and FILLSHRT) performed much of the remaining upper-air data preprocessing. The first program copied soundings at the beginning and end of the year, when necessary, to ensure that upper-air data bracketed the entire calendar year of surface data, as required by Calmet. Since Calmet requires the surface level and top level to be present and complete for all soundings, the program either filled in the missing data or output informative messages indicating where manual substitution was required. Following EPA guidance^{7,8}, missing data at the top or surface level were interpolated in time from the same station's data for one or two consecutive missing observations and were substituted from a nearby station's data at the same time for three or more consecutive missing observations.

Unlike the requirements for surface meteorological data, Calmet requires soundings to be present for every standard 12-hour observation time. Since no soundings may be left missing, periods with completely missing soundings were filled in either by

execution of the NDDH program FILLSHRT or manual substitution. For periods with one or two consecutive missing soundings, the program filled in the missing soundings from adjacent soundings in time for the same station. The substitutions were designed to retain appropriate representative diurnal variations in the substituted soundings. For periods with three or more consecutive missing soundings, the missing soundings were substituted by copying soundings from a representative nearby station for the same times using a text editor. Some editing of the substituted soundings was required to adjust the new soundings to the new station and its different elevation. Program FIXR62 also found missing temperatures, which were converted incorrectly in READ62, and replaced them with the correct code (in the correct units). The result was a file for each upper-air station and year containing sounding data in Calmet-ready format for every 12-hour observation.

The NDDH procedure for preparation of upper-air meteorological data is depicted schematically in Figure 2-4.

2.4 Precipitation Data

Hourly precipitation data for the five-year period 1990-1994 were obtained from Earth Info, Incorporated (Boulder, CO). Data were included for 96 stations located in North Dakota, eastern Montana, northern South Dakota, and western Minnesota. Location of these stations is shown in Figure 2-5.

Software provided with the Earth Info distribution allowed extraction of hourly precipitation data in TD-3240 variable record length format. The Earth Tech program for processing precipitation data (PXTRACT) requires data in TD-3240 fixed record length format. Therefore, the NDDH prepared a program (CONV3240) to convert precipitation files from variable to fixed record length format.

Earth Tech program PXTRACT was executed to extract individual station precipitation data from the TD-3240 files, and PMERGE was executed to consolidate individual station data into the single file required by Calmet (PRECIP.DAT). No substitutions were made for missing data (i.e., Calmet substitutes internally from nearest available station).

2.5 Geophysical Data

Most of the terrain elevation and land use data required by Calmet were downloaded from the United States Geological Survey (USGS) internet web site. Grid cell terrain elevations were derived from 1:250,000-Scale Digital Elevation Models (DEM) and land use data were derived from 1:250,000-Scale Land Use and Land Cover (LULC)

Figure 2-4
Upper-Air Met Data
Preparation

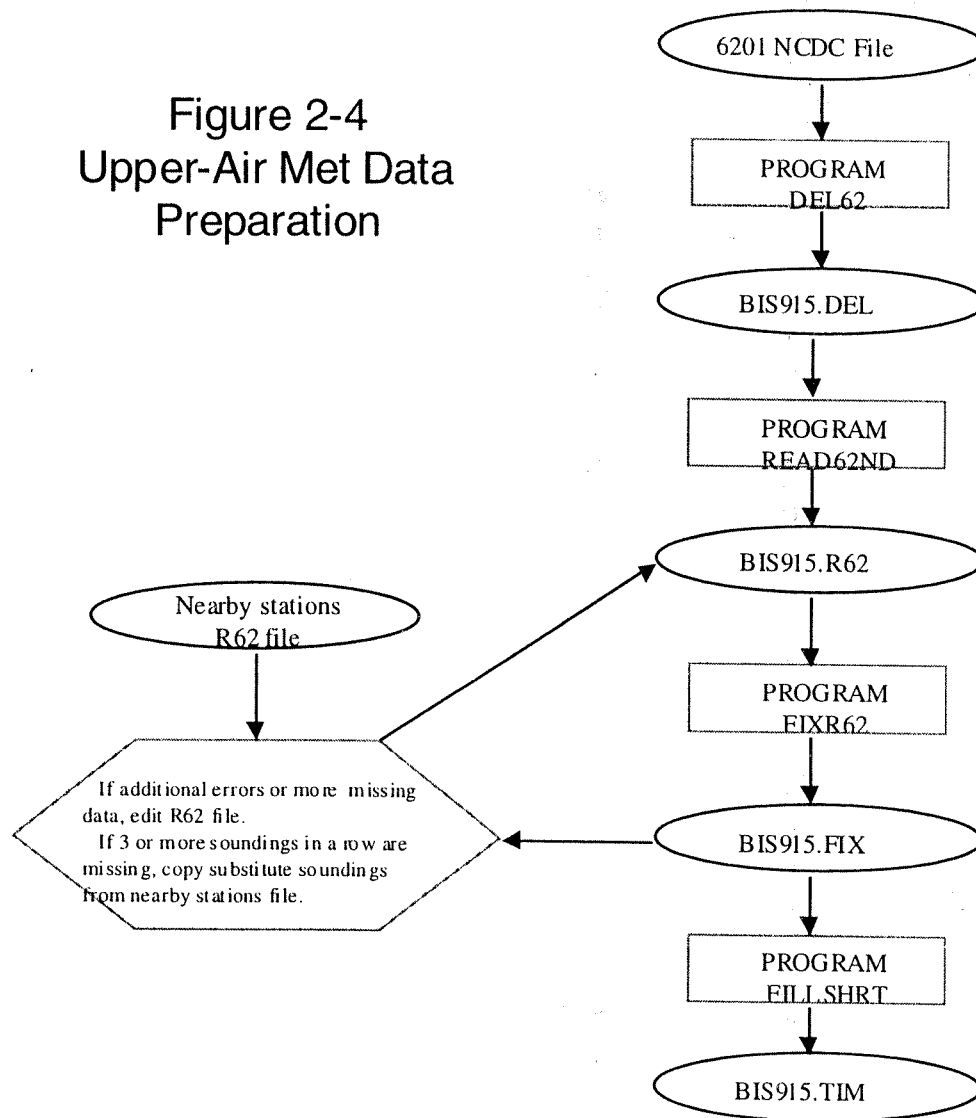
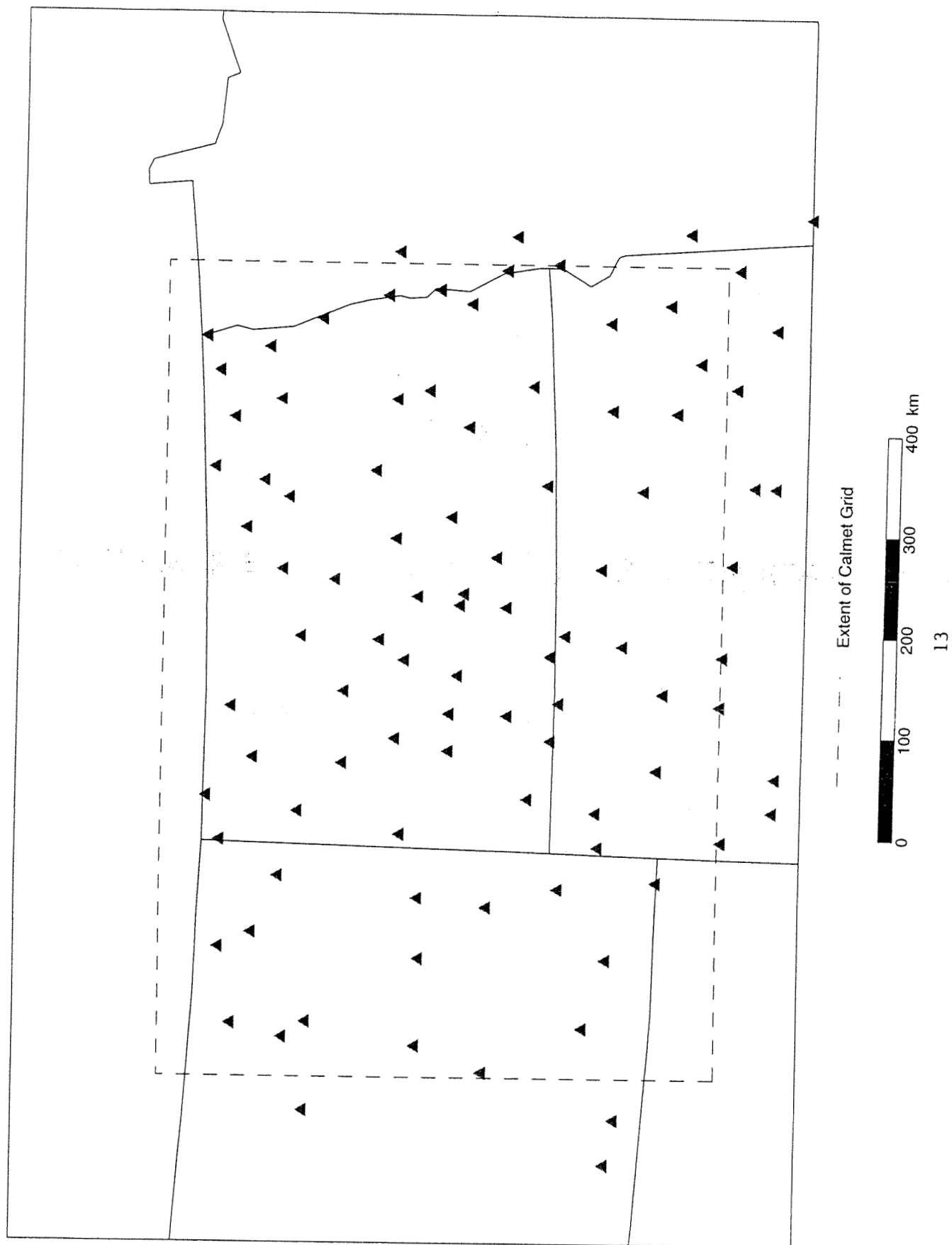


Figure 2-5: Precipitation Stations



files. The LULC coding of land use categories is equivalent to that required by Calmet.

The NDDH Calmet grid extends into Canada, which is not covered by USGS DEM's or LULC files. Terrain elevations for Canada were obtained from DEM's included with the EPA CD distribution "Calmet, Calpuff, and Calpost Modeling System Version 1.0". Because of the relative homogeneity of land use in northeastern Montana and northern North Dakota, land use categories for grid cells located in Canada were simply extrapolated from adjoining Montana and North Dakota grid cells.

Current Earth Tech programs for processing terrain elevation and land use data, and creating the GEO.DAT file required by Calmet were not available at the time the NDDH developed geophysical data. For terrain elevation processing, the NDDH developed software to allocate DEM data values to appropriate Calmet grid cells (TERSUM), and then average over all allocated values to determine the composite elevation for each cell (TERAVG).

Land use processing commenced with allocating LULC data values to corresponding Calmet grid cells (via NDDH program TERSUM). For each grid cell, a composite land use type was assigned based on the most frequently allocated value. Other land-use-related parameters required for the GEO.DAT file include surface roughness length (Z_0), albedo, Bowen ratio, soil heat flux, and leaf area index. Values for these parameters are related to land use type as shown in the Calmet User's Guide⁵. The distribution of land use types within each grid cell was used to establish composite values for these parameters. For example, if half of the LULC land use data values allocated to a grid cell were type 10 (Bowen ratio = 1.5) and the other half type 20 (Bowen ratio = 1.0), the composite Bowen ratio for the cell was set to 1.25. In the case of Z_0 , the composite value reflects the log-weighted average. NDDH coded the above procedure into program TERAVG. Coding was consistent with the CMLAND program provided by John Vimont of the National Park Service, Denver, CO (NPS).

The resultant fields of terrain elevation, land use type, Z_0 , albedo, Bowen ratio, soil heat flux, and leaf area index were edited into the GEO.DAT file for processing with Calmet.

2.6 Calmet Testing

Once the requisite surface, upper-air, precipitation, and geophysical data files were prepared, Calmet was extensively tested to determine optimum settings for control file options and

parameters. John Vimont (NPS) provided initial advice on control file settings⁹.

For testing purposes, Calmet was modified to optionally output Surfer-compatible XYZ files¹⁰ for winds (all levels), stability class, and mixing height for the entire Calmet grid for a selected time frame, in order to plot the horizontal distribution of these variables to better judge the appropriateness of Calmet's processing. A Surfer script was prepared to "mass produce" hourly plots of these three parameters for the selected time frame (usually 24 to 48 hours).

The NDDH examined several episodes of plotted wind vectors, stability classes, and contoured mixing heights, with emphasis on episodes (1990-1994 data) where winds might direct significant source emissions toward Class I areas. Episodes included cases with frontal passage or other wind shifts. During the iterative testing process, Calmet control file settings were individually and systematically adjusted primarily for wind and mixing height parameters. Parameters were adjusted so that plotted fields converged to a realistic and relatively smooth appearance. Output wind fields were examined to ensure that spatial variations due to frontal passage and terrain effects were reasonable, and to ensure a realistic transition from surface through upper-level winds. This testing process resulted in the final control file settings used with Calmet, as summarized in Section 2.7.

One problematic issue which arose during the testing of Calmet was a chronic discontinuity between surface and upper wind levels. To mitigate this problem, the option to extrapolate surface wind observations to upper layers was deployed, using similarity theory (Option 4) and layer-dependent biases. Calmet Version 5 extrapolates surface winds both for setting the initial guess field, and for introducing observations in the Step 2 wind field. Unfortunately, the model utilizes the bias factors for the initial guess field, only. The Step 2 vertical extrapolation has equal effect through all upper layers. The NDDH felt this was unrealistic because resultant upper layer wind fields reflected anomalous surface-layer (low-level) perturbations consistently, upward through all upper layers, even in the top layer (4000m). It was felt that such low-level features should dampen with height and not extend up into the middle troposphere. In other words, the Step 2 vertical extrapolation essentially undid the effective Step 1 (dampened) vertical extrapolation of the wind fields. Therefore, the NDDH modified the Calmet code to simply eliminate the vertical extrapolation in Step 2, resulting in a more realistic transition from surface to upper layers. NDDH changes to Calmet code are documented in Appendix A.

2.7 Calmet Execution

Calmet was executed with surface data, upper-air data, precipitation data, and geophysical data as described previously, and with control file options/parameters established in the testing process outlined in Section 2.6. More significant control file settings are summarized in Table 2-1. The complete NDDH Calmet input control file is provided on computer media with this report.

Calmet was executed individually for each year of meteorological data (1990-1994). Resultant processed files were used for Calpuff modeling of MRY station impact on Class I PSD increments and visibility. The processed file for each year required about 850 megabytes of disk storage.

Table 2-1
Calmet Control File

<u>Parameter/Option</u>	<u>Value</u>
No. surface stations	24
No. upper-air stations	6
No. precip stations	96
No. X grid cells	40
No. Y grid cells	27
No. vertical layers	8
Diagnostic wind module	Yes
Use O'Brien procedure	No
Extrapolate surface wind observations	-4
RMAX1	300 km
RMAX2	1200 km
TERRAD	100 km
R1	125 km
R2	100 km
No. barriers (NBAR)	0
MNMDAV	5
ILEVZI	4
Minimum overland mixing height	20 m
Maximum overland mixing height	4000 m
TRADKM	300 km
SIGMAP	100 km